

KurokoPointer: プレゼンテーションにおける 確実なポインティングのための話者映像処理

梅村 恭司^{†1} 梅村 真由^{†2}
山根 恵和^{†1} 大園 忠親^{†2} 新谷 虎松^{†2}

概要: スライドを用いたプレゼンテーションにおいて、説明個所の確実な指し示し(ポインティング)は重要であるが、聴衆のすべてが示した場所を認識するまで、場所指定を継続して維持することは努力を要する。本稿では、自然で確実なポインティングを実現するシステムについて述べる。このシステムはコンピュータにあるカメラを利用し、話者のシルエットを生成する。その上で、手に持った目立つ物体を追跡して、シルエットと物体の場所を示すマーク(ポインタ)をプレゼンテーションのスライドの上に重ね合わせて表示する。物体の動きが停止したことを検出すると、別の形状のマーク(スタンプ)を生成し、それを聴衆にとって十分な時間だけ保持する。このとき、物体が動いたとしてもスタンプは移動しない。このようにして確実なポインティングを実現した。

KurokoPointer: A Reliable Method for Pointing at Presentation Contents by Processing Presenter's Image

Kyoji UMEMURA^{†1} Mayu UMEMURA^{†2}
Yoshikazu YAMANE^{†1} Tadachika OOZONO^{†2} Toramatsu SHINTANI^{†2}

Abstract: Though pointing devices in presentation play an important role in presentation, it is usually difficult to keep the pointing action long enough so that all audience should recognize the place to be pointed at. In this paper, we present a presentation system that makes reliable pointing possible. This system utilizes the camera on computers and captures the silhouette of the presenter, and then generate pointing image in slides by recognizing a salient object in the hand of the presenter. When the object stops, the system responds to this event and then generates a different kind of pointing image and keeps the image long enough for audience, even if the salient object moves to another place. We have, thus, realized a reliable method to point at the places on presentation slides.

1. はじめに

オンラインでプレゼンテーションが配布されるという状況は多くなっているが、そのときにプレゼンテーションスライドを固定カメラでストリーミングするという状況で、話者のシルエットをポインタとして利用するという試みがなされている[3]。同じ状況のために、深度画像が取得できるカメラで、機能追加されたシステムを作成されている[4]。この先行研究では、被験者実験によってプレゼンテーションの効果を測定しているが、話者シルエットの利用だけでなく、話者の動作によって、ポインティングしている場所にマーカを生成し、それを一定時間保持することが効果的であることが示された。文献[4]において、プレゼンテーションの印象がポインティング機能で有意に改善できている要素は、指し示した場所が聴衆に十分な時間の長さで保持

されていることが示唆されている。しかしながら、このシステムの方法では、ユーザの動作の認識が基盤にあるため、実際に使うには専用のカメラを用意しなければならないという問題があった。

また、深度画像に加え、指の動作の検出のための専用のデバイスも追加し、プレゼンテーションシステムの操作性が補強されたものが開発されている[5]。この報告において、特殊デバイスを使うことの問題が示され、通常のカメラを用いてポインティングを行うことについての機能も示されている。それは、話者シルエットと特徴的な色をもつ物体を用いてポインティングする方法である。

このような背景から、カメラを利用するが、深度情報も、指動作の検出のためのデバイスも用いない条件で、文献[4]で提案された機能：保持機能のもつポインティングデバイスの実装方法を報告するのが、本稿である。本稿の内容は、文献[3]と比べて、ポインティングの機能の強化がなされ、文献[4]と比べて、聴衆に提供する機能は同一だが、話者にとってはシステムの利便性(特殊カメラを必要としないこと)を提供している。文献[5]と比べると、確実なポインティングのための場所の保持機能が異なる。

^{†1} 豊橋技術科学大学
Toyohashi University of Technology

^{†2} 名古屋工業大学
Nagoya Institute of Technology

2. ポインタとスタンプ

文献[4]において、ポインティングを確実に行う工夫として、スタンプ機能が提案されている。文献[5]で示されるポインティング方法も、通常のポインティングデバイスも、場所の保持の機能がないため、話者が場所を指し示したときに聴衆がそこを見ていないと、指そうとしていた場所が聴衆に伝わらないままにポインタが移動してしまうという危険性がある。また、実際にプレゼンテーションでは、そのようなことが生じる。

文献[4]におけるスタンプ機能とは、話者が指定した場所に円形のマークを表示し、それを20秒保持し、そのあと10秒かけて徐々に消すというものを、シルエットによるポインティング機能とともに実装したものである。ポインタが動いても、スタンプは20秒の間、動かないで表示されるために確実に指示している場所が聴衆に伝わる。

文献[4]では下記の条件で印象の比較を行っている。A: 話者のシルエットとポインタとスタンプの表示, B: 話者のシルエットとポインタ表示, C: 話者のシルエットのみの表示, の3つの条件である。そして、14人の被験者に対し、15分を区切りとして条件を変更して、3条件でプレゼンテーションをして、印象の比較で符号検定を行ったところ、AとB, AとCの間の印象の差があるあり、いずれもAの条件のほうの印象が良いと答えた人数比が13:1で、差は危険率1%で有意と報告されている。

3. シルエット生成方法の変更

文献[4]で、シルエットがちらつくとプレゼンテーションの印象が低下することが報告されている。通常のカメラでのシルエットの利用を行っている文献[3]でのシルエットの生成方法は、動きのあるところをシルエットとし、ポイントしているところを目立たせる設計を行っていた。動きのあるところを表示すると、必然的にシルエットがちらつくことになる。文献[3]では、指し示す場所を強調するために動いているところをシルエットとしていたが、指し示している場所を別の方法で示すことにすると、ちらつくという欠点しか残らない。このため、本稿ではシルエットは動きがあるところにせず、背景との差分があるところとする方針にした。図1にシルエットを持たないオリジナルのスライドを示す。図2は文献[3]のシルエット像を図1に重ねて影を表示したものである。手が動いているので、手の部分に影ができていのが分かる。図3が本稿で使用したシルエット像である。背景と差があるところが影として表示されている。文献[3]のシステムでは、動いている部分の輪郭部分が影となっており、全体的に指先が強調される。

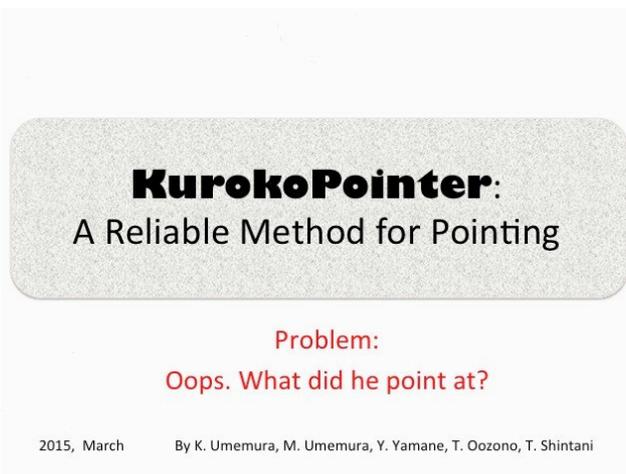


図1 例で用いるスライド
Figure 1 Slide used for examples

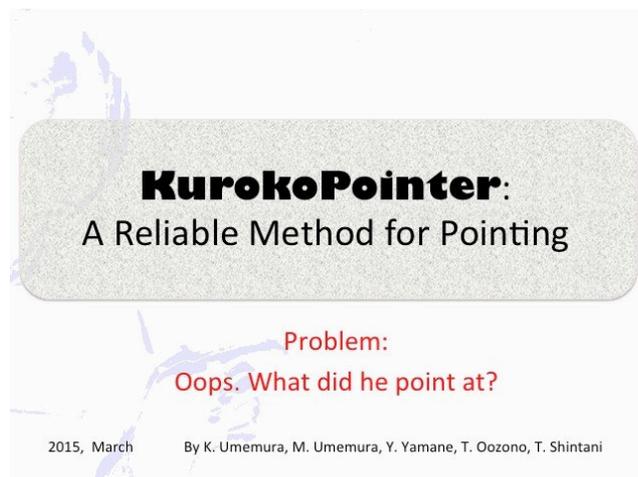


図2 文献[3]におけるシルエット像
Figure 2 Silhouette described in reference [3]

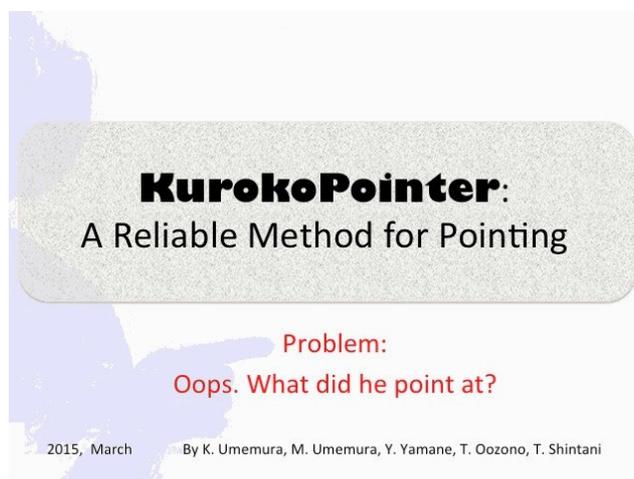


図3 背景差分のシルエット像
Figure 3 Silhouette by difference from background

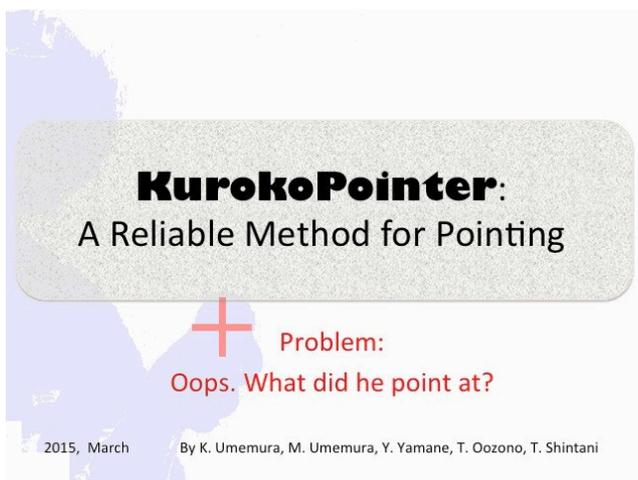


図 4 ポインタの場所が十字で示されている
Figure 4 Red cross mark indicates the current pointer.

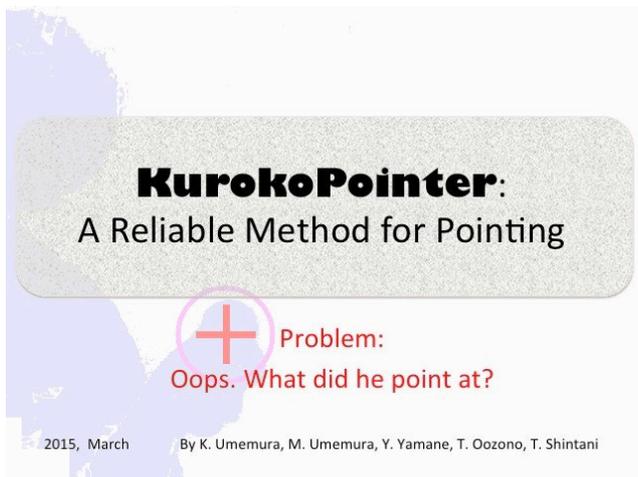


図 5 ポインタが静止することでスタンプが現れる
Figure 5 Stamp mark is generated where the pointer stays.

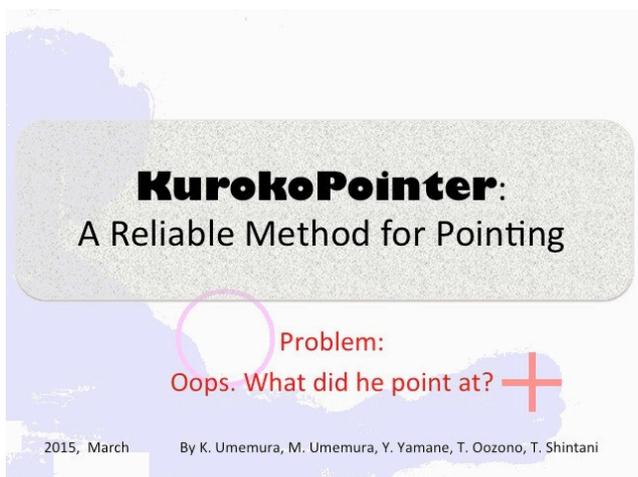


図 6 ポインタが動いてもスタンプは動かない
Figure 6 Stamp mark remains the same place.

今回のシステムでは、文献[4]のシルエットと同様、輪郭で囲まれた部分が塗りつぶされた像を使うため、影の部分が多くなり、影の動きはちらつきが少なくなるが、注目する部分（手）が分かりにくい状態になっている。

4. スタンプ機能の実装方法

文献[4]では、話者が手をカメラにむかって押し出すとスタンプが表示される設計となっていた。通常のカメラだけでは、この動作を確実に検出することは難しい。そこで、話者にほかの場所と区別できるような目立つ色の物体を手にもってもらい、それが一定時間動かないことを検出して、スタンプを生成することにした。

文献[4]と類似動作、すなわち、物体をカメラ方向に押し出すことによる動作によって、重心が動くことなく物体領域が増加することを検出して、スタンプを生成することも検討したが、重心を保持した押し出しは訓練が必要な動作であり、確実にスタンプを生成できなかった。

一方で、物体が停止することによるスタンプの生成は、意図しないスタンプが生成されることが危惧されたが、手に保持したものを動かさないのは、はっきりとした意図が必要であり、実際には誤動作が起きにくかった。このように、スタンプを生成する話者の操作は文献[4]と異なるものとなった。

5. システムの動作

システムは、jpegのスライドファイルを用意し、それを用いてプレゼンテーションをするものである。Power Pointそのものを使用するものでないが、文献[5]などの実装をすれば、そのまま使うシステムを作成する方法もある。本システムの特徴は、コンテンツの背景（通常は白の部分）に話者映像を処理したものを重ね合わせ、ポインティング機能を実現しているところである。

図4、図5、図6によって、ポインティング動作の説明をする。図4は手に目立つ色の物体をもっているときに表示される影をスライドに重ねたものである。この例では物体は赤い玉で、赤い画素の領域の重心が十字の中央になるように表示される。図4は、図3と比較すると、注目すべき箇所が明確になっている。

図4のポインタは、手の動きに追従する物であるが、このポインタを1秒程度静止させると図5の状態になる。完全に静止していない場合でも、移動速度がある閾値より小さい状態が1秒弱続いた場合には図5の状態になる。実際には、重要なところを精密に指定しようとして狙いを定めると、狙いが定まったところに円が出現するように閾値を調

整した。

スタンプが出現したあとに手を移動した状態が図6である。十字のポインタは手に追従して移動するが、スタンプのマークは生成された場所から動かない。この状態が20秒あり、その後、10秒の時間を使って円が薄くなり、見えなくなる。この時間をとることによって確実なポインティングができています。

スタンプが消える前に、別の場所にポインタを静止させてスタンプを生成すると、その時点で古いスタンプは消える。このようにしたこと、ひとつの理由は、複数の場所に注目点が存在すると混乱することであり、もうひとつの理由は、スタンプの場所がずれていた場合に修正することが自然にできるように考えたことである。また、スタンプが消えるまえにスライドを移動した場合は、スタンプは消える。これは、注目点はスライドに属しているため、スライドが消えたら、注目点はなくなるため、それに同期してスタンプが消えるのが妥当と考えた。

ここで、ポインタとスタンプの生成時点において、シルエット人物像がいることにより、ポインタとスタンプがスライドのコンテンツと別のレイヤの情報とわかる。これが特徴と考え、システムを KurokoPointer と呼ぶことにした。

6. システムの有用性

まず、ポインティングの考え方については文献[4]に準じており、システム利用者の印象もほぼ同じであり、聴衆によっては、同じシステムと勘違いすることもあった。しかしながら、本稿の方法で利用するデバイスは、通常のカメラと目立つ色の物体であり、特殊カメラを必要としないため、文献[4]との比較では有用性がある。

一方で、文献[3]と文献[5]は、特殊カメラを使用しない件でのシステムの提案があり、その意味では本稿と同じ条件を想定しているものである。しかしながら、ポインティングのために提供する機能が異なる。その比較には文献[4]における、ユーザ評価の結果を使うことができる、条件A（話者のシルエットとポインタとスタンプの表示）が、本稿のシステムに相当する条件であり、条件B（話者シルエットとポインタの表示）は、文献[3]における特殊デバイスを使用しない条件でのシステムに相当する条件であり、条件C（話者シルエットの表示）は文献1に相当する条件と考えられる。相当する条件で、差が報告されているので、本稿の提案は、文献[3]、および、文献[5]における通常カメラを利用する条件での提案に比べて有用性がある。

7. シルエットについて

提案するシステムの着想の出発点として、協同作業空間の実現における人物映像の利用の報告[1]がある。共同作業空間で、共同作業者のプレゼンスを協同作業空間に重ねると

いうアイディアである。これをプレゼンテーションに利用しようとしたとき、人物映像とプレゼンテーションコンテンツを重ね合わせることによる干渉が問題と考えたため、人物の存在をシルエットにして、コンテンツを重ねたとしても、コンテンツを読み取る事を邪魔しないようにすることを考えた。シルエットを用いることで、これを用いないときよりもプレゼンテーションに効果があるという結果が文献[3]に示されているため、存在感を伝えることに一定の効果を示唆される。

一方で、シルエットにすることにより、表情などが失われる。表情をもつイメージとコンテンツを重ね合わせるために、コンテンツを変形させるという提案がある[2]。これを可能とするには、プレゼンテーションのコンテンツに対しての条件が必要であり、それが満たすとシルエットよりも聴衆に対して話者に関する多くの情報を伝え得るが、コンテンツの準備のコストがあると考えられる。

最後に、ポインタ、スタンプなどを実装したので、シルエットは不要ではないかという意見もあるが、シルエットがあることによって、コンテンツとポインタが明確に分離すること、話者が動作していることが了解できることについて、聴衆の注意の継続に益すると考えているが、これについての評価/評価方法について、今後考えていきたい。

8. おわりに

本稿では、通常のカメラと色のついた物体を用いて、話者が指定する場所を確実に聴衆に伝えるためのシステムについて述べた。提案方法の特徴は、動き回るポインタとともに、一定時間動かない別のマーク（スタンプ）を表示し、それを十分に長く表示することである。実現した機能は、既存の研究[4]と同じであるが、特別なカメラを利用せず、カメラ付きのノートコンピュータであれば使用できる。

謝辞 本研究は JSPS 科研費 26330396 の助成を受けた。

参考文献

- 1) 小林 稔, 石井 裕:「ClearBoard-2における協同作業空間と会話空間のシームレスな融合」, 情報処理学会, 情報処理学会研究報告, [グループウェア] 93(34), pp43-50, 1993
- 2) 前田 晴己, 黒澤 祐也, 栗原 一貴, 宮下 芳明:「MAEDE: スクリーン前でのプレゼンテーションスタイル」, WISS2011, 第 19 回インタラクティブシステムとソフトウェアに関するワークショップ 論文集, pp. 164-166, 2011.
- 3) 梅村 恭司, 梅村 真由:「Kuroko: 話者シルエットを活用するプレゼンツール」, 情報処理学会インタラクティブ 2012, pp.281-286, (2012)
- 4) 若松 翔, 梅村 恭司, 三輪 多恵子, 岡部 正幸:「話者動作による確実なポインティングのできるプレゼンテーションシステム」, 情報処理学会インタラクティブ 2013, pp.750-755, (2013)
- 5) 山田 裕之, 丹羽 佑輔, 白松 俊, 大園 忠親, 新谷 虎松:「即応的プレゼンテーションのための PowerPoint スライドオブジェクト操作機構の実現」 信学技報, Vol. 113, No. 332, pp. 59-64, 2013.